



Oy	303	ALKEVYSTOTKOLFEKIKOGTVDFENKEIROPKSALKLEVSDANDLEFNKT-----G	386
		:   :   :   :   :   :   :	
Dd	323	AVTKLVPREHQAIINVVGSTKHIAN-GVROLSRIKE----PSQOIIREKLPHYLSKA	377
Oy	357	OCTVEINNEIRD--PSKALLIRKYSTGAEADLFENKIGOGTVDFEINNEIRDPSALLRKUY	414
		:  :  :  :  :  :  :  :  :	
Dd	378	KCAVEHVAKKVASVYPIK---ORGDOPSENAVETVPSG--BSAETFEVEBEQOYDAVT	432
Oy	415	T-----EADD	419
		:   :   :   :	
Dd	433	TOEVNSEKVADD	445

## RESULT 2

merozoite surface antigen, 60K - Babesia bovis  
C:Species: Babesia bovis  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 17-Nov-2000  
C:Accession: S27778  
R:Suarez, C.E.; Palmer, G.H.; Jasmer, D.P.; Hines, S.A.; Perryman, L.E.; McElwain, T.F.  
submitted to the EMBL Data Library, February 1991  
A:Description: Characterization of the gene encoding a 60 kilodalton Babesia bovis merozoite surface antigen  
A:Reference number: S27778  
A:Accession: S27778  
A:Molecule type: mRNA  
A:Residues: 1-565 <SUN>  
A:Cross-references: EMBL:M38218; NID:q155883; PID:q155884  
C:Superfamily: merozoite surface antigen p58  
C:Keywords: surface antigen

Query Match	35.08	Score	826.5	DB 2	Length	565
Best Local Similarity	35.58	Pred. No.	2.3e-46			
Matches 178; Conservative	79	Mismatches	195		Indels	49
					Gaps	6

```

0Y 1 MAPSOSGVDTTLLAASSVDSAAVAAYMINSPMSDYLSAVSNGFMRICSOYPKSNC 60
Db 31 LAPAEVGDJLTLEADPLMTLRDHMHNTKDMKAVLSNGRQIYNDVCSNAPEDSNCR 90
0Y 61 ASVASVMSFCANODCLTLOSLKYPLEAKYOPLTLPDPYOLEAFILFEKESDANPANSTER 130
Db 91 EYVNNYADCEMGEYGFETIDNVKXYPLEOYOPLSLPYOLDAAFRLFEKESAPKANSVK 150
0Y 121 RFWMRFRRCKNHSYFADVLVFNLEKVNTRDAOATDLENFASRLYMAATLYKTYNTNDEF 180
Db 151 RRMFLMFRGANCNGDHYEFTYGLLNNNVNHEEGTIDVEYLNVKKLYATMYTKTYLYNSM 210
0Y 181 GASFNKLSFTTGLFGWGIGIKALKQIIRSNLPJLDIGTEHSVSLQIHTSSYKDYMDTOIP 240
Db 211 NAKFENRFEFTTKTIFSRRIQOTLSRIIRNNVPEDF-EERSIERITOLTSSYEDYMLTOIP 269
0Y 241 ALPFKAKRSLWVORLLATVAGYVDPWYKMKYMLKMPNVRVILPIPKF----- 282
Db 270 TUSKARRADWKKVLLGSLTSYEAAPWKRRIKKFRDEFSNANVOPTKKFLEDINEVT 329
0Y 293 -----FNKEIRBPSKALKEKYSTDKLFEKNGIGOTVD 336
Db 330 KNYLKANVAEPTKPEMODTNEHTKGYLKENVAEPTKFEKAPQVYKHFEDENIGOPTKE 389
0Y 327 FPNKKEIRDSKALKKEKVSNDADLFENKLTGCGYVDPINNEIRBPSKALLRKVSTGAEJLF 386
Db 390 FFERAPQATKHFLENIGOPTKEFF-REAPQATKHFLENIGOPTKEFFRDVYQVTKKVL 448
0Y 387 ENKIGOGTVDF-----INNEIRBPSKALLRKVYTEADDLEFNKIGOGTVDFINK 435
Db 449 TENINAOPTKEFREVRPHATMKVLNENIADQPAKEIIEHFPETGAKN-FISAHBETKQPLNE 507
0Y 436 EIRBPSKALLIR-KVSTEADNL 455
Db 508 TVGQPTKEFLNGALETTKDAL 528

```

B45561  
merozoite surface antigen (clone 7) - Babesia bigemina  
C:Species: Babesia bigemina  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 16-Feb-1997  
C:Accession: B45561  
R:Mishra, V.S.; McElwain, T.F.; Dame, J.B.; Stephens, E.B.  
MOL. Biochem. Parasitol. 53, 149-158, 1992  
A:Title: Isolation, sequence and differential expression of the p58 gene family of Babesia  
A:Reference number: A45561; MUID:92365724  
A:Accession: B45561  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-442 <MIS>  
A:Cross-references: GB:M85184  
C:Superfamily: merozoite surface antigen p58  
C:Keywords: surface antigen

Query Match	34.98;	Score 823;	DB 2;	Length 442;
-------------	--------	------------	-------	-------------

Best Local Similarity 42.3%; Pred. No. 2.9e-46;  
Matches 174; Conservative 78; Mismatches 129; Indels 30; Gaps 10;

QY 4 SDSVGDVTKTLAASESYDSANAYMINSDMDYLSAVSDNFAERICSQVPKGSNCASV 63

```
Db 33 AEVGDVSKTLLLEANEVVNMEATQYKMDQSQSNVKETIVGEVCEKVAAGNSTGSEV 92
```

QY 64 SAYMSRCAGQDCDLTLQSLKYPLEAKYQBLTLPDPYQLEAFILFKESDANPANSTEKREWF 123

```

Db      93 IAVVNRCEDEGCLTLD SM-----KYPRLSLPNPYQDDA FMLFRESDSNPAKNEVKRFW 146

```

124 MRFRRGNHSYFHDLVFNLLKKNVTRDADATDIENFASRYLYMATLYYKTTYTNDEFGAS 183

```

Db 147 MBSR--SSHGDYHHEFVSLIKKNVRDEESNDVENEASQYFYMTLLYKTYLTFVDTAAK 204

```

184 FNNKLSFTTGIFGCGIKRAIKQIIRSNNPLDIGTEHSVSRLOHTTSYKDYMDQIPALP 243

```
Db      205 FENKLAFTRLFGEGCIQKALKRVLRSNLPVDLGT-HPEATIREIASGCEGYMMTOVPAMT 263
```

QY 244 KPAKRESLMVQRLLATVAGYVDTPTWKKKWYMKLNEMVNRVFIPTKKEFNKEIREPSK - 302

Db 264 SFAERFSKMATKTLVTVSDYVHLPAYKRWYRKFEKIVN-FTDPAKLIMKHVSQPKV 322

303 ALKEKYSTDTKDLFENKIGQTVDFEFNKEIRDPSSALKEKVSNDAKDLFENKI-----G 356

Db 323 AYTKLVP EEHQAIRNVGSGSTKHIAN-GVRLDARIKE---PSQIIR EKLPHYLSKA 377

357 ÖGTVEFINNERDPSKALIRKVGÆDLFEN-----KIGÖTVEFINNE 401

## RESULT 4

D45561  
merozoite surface antigen (clone 14) - *Babesia bigemina*

C:Species: Babesia bigemina  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 16-Feb-1997

C:Accession: D45561  
R: Mishra, V.S.; McElwain, T.F.; Dame, J.B.; Stephens, E.B.

Mol. Biochem. Parasitol. 53, 149-158, 1992

A:Reference number: A45561; MUID:92365724  
A:Accession: D45561

A: Status: preliminary  
A: Molecule type: DNA

A:Residues: 1-480 <MIS>  
A:Cross-references: GB:M85187

C:Superfamily: merozoite surface antigen p58  
C:Keywords: surface antigen

Query Match	34.58;	Score 813.5;	DB 2;	Length 480;
-------------	--------	--------------	-------	-------------

Best Local Similarity 41.68; Pred. No. 1.3e-45;  
Matches 180; Conservative 74; Mismatches 142; Indels 37; Gaps 12;

4 SDSVGDVTKTLAASEVSDSAANNYMINSDMSDYLSAVSDNFAERICSQVPGKNCASV 63

```
Db 33 AEVVDGVSSTLLAANEVVAEMEQAQINEDMKIQLANVKEITVDEVCGRDAGSPICRSKV 92
Oy 64 SAYMRCACODCLTLOSILKYPLEAKYQPLTPDPYOLEAFLTFKESDANPANSTKRFW 123
Db 93 IAYVRCDEGDCDLTLDMS-----KYKPLSLPNPYOLDAAFLMFRSDSNPAKNEVKRFW 146
Oy 124 MFRGRKHNHSYFHDLVFNLEKNVTARDADATDIENFASRYLTMATLYXTYTNVDFGAS 183
Db 147 MRSR--SSHGDIHFHVSVSLKKNVARDPSNDVENFASQYFMTLLYXTYTLVDFTAAK 204
Oy 184 FENKLSFTTGLFGWGIKRALKOIIRSNPLDIDGTEHSVRLQIHITSSYKDYMDTOIPLP 243
Db 205 FENKLAFTTFLRFGIQRKRLRVSNLPLVDLGT--HPEATIREINAGCYEYMMTQVPAMT 263
Oy 244 KRAKRFSLMVOORLATYAGYDTPWKYKWKYKLNFMVNRVFIPTKRFKNEIREPSK- 302
Db 264 SFAGRFPSKMATKTLTLYTSDVYHLPAKRYKRFKEFIYN--FFTDPAKLIMKHVSQPVKT 322
Oy 303 ALKEKYSTDTKDLFENKIGOGTVDFFNKEIRDPSKALKEKVSNDADKDLFENKI-----G 356
Db 323 ATKLVPEEHRAIRNVVGOSTKHIAN-GVRLSRMIKE----PSQOITREKLIPLYLSKA 377
Oy 357 OCTVDFINNEIRD--PSKALIRKVGSTGADLEFNKIGOGTVDFINNEIRDPSKALIRKY 414
Db 378 KGAHVHVKKVKSVPPIK---QKGPSPSAVAEETVPSG--DSAEETFEVEPEEQYVDANT 432
Oy 415 T-----EADD 419
Db 433 TQEVNSEKVDADD 445
```

## RESULT 5

merozoite surface antigen (clone 9) - Babesia bigemina  
C:Species: Babesia bigemina  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 16-Feb-1997  
R:Accession: C45561  
R:Mishra, V.S.; McElwain, T.F.; Dame, J.B.; Stephens, E.B.  
Mol. Biochem. Parasitol. 53, 149-158, 1992  
A:Title: Isolation, sequence and differential expression of the p58 gene family of Babes  
A:Reference number: A45561; MUID:92365724  
A:Accession: C45561  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1442 <MIS>  
A:Cross-references: GB:M85185  
C:Superfamily: merozoite surface antigen p58  
C:Keywords: surface antigen

```
Query Match 33.5%; Score 790; DB 2; Length 442;  
Best Local Similarity 41.4%; Pred. No. 4e-44;  
Matches 170; Conservative 78; Mismatches 133; Indels 30; Gaps 10;  
Oy 4 SDSVSDVDTKTLAASESVDSANAYMINSMDSYLSASVNDNAERICSOVPGKSNCSASY 63  
Db 33 AEVVDGVSSTLLAANEVVAEMEQAQINEDMKIQLANVKEITVDEVCGRDAGSPICRSKV 92  
Oy 64 SAYMRCACODCLTLOSILKYPLEAKYQPLTPDPYOLEAFLTFKESDANPANSTKRFW 123  
Db 93 IAYVRCDEGDCDLTLDMS-----KYKPLSLPNPYOLDAAFLMFRSDSNPAKNEVKRFW 146  
Oy 124 MFRGRKHNHSYFHDLVFNLEKNVTARDADATDIENFASRYLTMATLYXTYTNVDFGAS 183  
Db 147 MRSR--SSHGDIHFHVSVSLKKNVARDPSNDVENFASQYFMTLLYXTYTLVDFTAAK 204  
Oy 184 FENKLSFTTGLFGWGIKRALKOIIRSNPLDIDGTEHSVRLQIHITSSYKDYMDTOIPLP 243  
Db 205 FENKLAFTTFLRFGIQRKRLRVSNLPLVDLGT--HPEATIREINAGCYEYMMTQVPAMT 263  
Oy 244 KRAKRFSLMVOORLATYAGYDTPWKYKWKYKLNFMVNRVFIPTKRFKNEIREPSK- 302  
Db 264 SFAGRFPSKMATKTLTLYTSDVYHLPAKRYKRFKEFIYN--FFTDPAKLIMKHVSQPVKT 322
```

```
Oy 303 ALKEKYSTDTKDLFENKIGOGTVDFFNKEIRDPSKALKEKVSNDADKDLFENKI-----G 356  
Db 323 ATKLVPEEHRAIRNVVGOSTKHIAN-GVRLSRMIKE----PSQOITREKLIPLYLSKA 377  
Oy 357 OCTVDFINNEIRDPSKALIRKVGSTGADLEFN-----KIGOGTVFINNE 401  
Db 378 KGAHVHVKKV--SKTFKRRAGESESESESTROSEEBELIKESQNSDRENDE 426
```

## RESULT 6

rhoptry protein homolog Bc60.2 - Babesia canis  
C:Species: Babesia canis  
C:Date: 01-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
C:Accession: C48572  
R:Dalympde, B.P.; Casu, R.E.; Peters, J.M.; Dimmock, C.M.; Gale, K.R.; Boese, R.; W  
Mol. Biochem. Parasitol. 57, 181-192, 1993  
A:Title: Characterisation of a family of multi-copy genes encoding rhoptry protein ho  
A:Reference number: A48572; MUID:93165069  
A:Accession: C48572  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1456 <DAL>  
A:Cross-references: GB:M91168; NID:9155908; PIDN:AAA27807.1; PID:9155909  
A:Note: sequence extracted from NCBI backbone (NCBIN:125083, NCBI:125084)  
C:Superfamily: merozoite surface antigen p58

```
Query Match 32.7%; Score 770.5; DB 2; Length 456;  
Best Local Similarity 35.0%; Pred. No. 7.8e-43;  
Matches 161; Conservative 95; Mismatches 161; Indels 43; Gaps 9;
```

```
Oy 1 MAPSDVGDTKTLAASSVDSANAYMINSMDSYLSASVNDNAERICSOVPGKSNCS 60  
Db 31 LSKSDGAEXKTLSTLLNVADSTRALEGRWNAAMANSNGRREDEEAVCGNIAETEGC 90  
Oy 61 ASVASMRCACODCLTLOSILKYPLEAKYQPLTPDPYOLEAFLTFKESDANPANSTK 120  
Db 91 KSAVEYVESCVRKDCSIEKQTPKEKEYOPLTPDPYOLEAFLTFKESDANPANSTK 150  
Oy 121 RFWMRFRGRKHNHSYFHDLVFNLEKNVTARDADATDIENFASRYLTMATLYXTYTNVDF 180  
Db 151 AFWMRFRGRGRCGAYHNFVNLITKLNLSMDVODNEGEVRYKAYVATMYKYKTYALDV 210  
Oy 181 GASFEKLSFTTGLFGWGIKRALKOIIRSNPLDIDGTEHSVRLQIHITSSYKDYMDTOI 240  
Db 211 NARIINKIAFSRHLFGRQIRNALNTIIRSNIPEDFG-KYNVDRLRHVMGGEYEMKQYV 269  
Oy 241 ALPKRAKRFSLMVOORLATYAGYDTPWKYKWKYKLNFMVNRVFIPTKRFKNEIREP 300  
Db 270 SLPNFKKTAGANVYKSLINNVAGAYQKOPWFKKLNNOIRNFVYKIHPTKEFFVYKIH 329  
Oy 301 SKALKKESVDTKDLFENKIGOGTVDFFNKEIRDPSKALKEKVSNDADKDLFENKIGOGTV 360  
Db 330 -----TKFEVKNKIHPTKEFFVYKIHPT-----KEFEVKNKIHPTK 367  
Oy 361 DFINNEIRDPSKALIRKVGSTGADLEFNKIGOGTVDFINNEIRDPSKALIRKYTTAD 420  
Db 368 EEFVNKLHEPTKEFFSNMVPAGAFKISEKAGR-----HLMS-SKYVPE--DEPSS 416  
Oy 421 FENKI--GOGTV-DFINKEIRD-----SKALIRKYTE 451  
Db 417 LENEAVEDQLTIMGDVTDEEMATPTYEQSGESLNEVGN 456
```

## RESULT 7

rhoptry protein homolog - Babesia ovis (fragment)  
C:Species: Babesia ovis  
C:Date: 01-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
C:Accession: A48572  
R:Dalympde, B.P.; Casu, R.E.; Peters, J.M.; Dimmock, C.M.; Gale, K.R.; Boese, R.; W  
Mol. Biochem. Parasitol. 57, 181-192, 1993



```

Query Match          5.7%: Score 134; DB 2; Length 1098;
Best Local Similarity 18.4%: Pred. No. 0.69;
Matches 61; Conservative 75; Mismatches 124; Indels 72; Gaps 13.

QY      167  ATLYKRTVNDVEFGASFNNKLSFTTGLFGWGIKKALKQIIINSNPLDIDGTEHSVRLQH 226
      ||:| |-----KIGELVAVI LAGKMWGCKMMLFAITTPFYKVLGIGIYVIGLKE 411
Db      361  ATIVYLT-----

QY      227  ITTSKAYMDIQIPLPFPFAKRFSLMV--VQRLATVAGYVDTPMYKKWYMKLKNPMV-- 282
      :|: |-----:|: |-----:|: |-----:|: |-----:|: |-----
Db      412  LAT---YVSSGF--LDFFKGAFFBIIIMWOKLIIMIS-----KAYTLKMLLVTFW 456
      :|: |-----:|: |-----:|: |-----:|: |-----:|: |-----

QY      283  -----NRPVPTKRF--FNKEIRBPSKALKESVSDTDLFENKIGIGG-- 323
      :|: |-----:|: |-----:|: |-----:|: |-----:|: |-----

Db      457  KKAEEKRDTSGSEPRDKKFDPNATININKKAAEDYOKLDDIEFNKRQRIY--NKTGKARE 515
      :|: |-----:|: |-----:|: |-----:|: |-----:|: |-----

QY      324  -----TVDFENKEIRBPSKALKESVSNDAKDLF-----ENKIGQTVDFINNEIRD 369
      ||: |-----:|: |-----:|: |-----:|: |-----:|: |-----

Db      516  QALRNLEKVTINENKKNQKPFDEYSKIFDQLDENKKLLVGYEKSVNPFNSNVFVNE--- 571
      :|: |-----:|: |-----:|: |-----:|: |-----:|: |-----

QY      370  PSKALIRKVTISGAEDLFENKIGQGVDF---INNEIRBPSKALIRKVTYEAADLFENKIG 426
      :|: |-----:|: |-----:|: |-----:|: |-----:|: |-----

Db      572  -YONLKKKEKESRERERIIITLPPTDQVSALQKLNDEINENKAFVEKYGKSFEYTL--NESN 628
      :|: |-----:|: |-----:|: |-----:|: |-----:|: |-----

QY      427  QGTVDFINKEIRBPSKALIRKVTSTADNLLK 458
      :|: |-----:|: |-----:|: |-----:|: |-----:|: |-----

Db      629  KOVVALEKOVNEYEKTLADRSFVPAQKALOK 660
      :|: |-----:|: |-----:|: |-----:|: |-----:|: |-----

```

RESULT 11  
PC6003  
surface membrane protein Imp4 - Mycoplasma hominis (fragment)  
N:Alternate names: hypothetical 624 protein; Imp4 protein  
C:Species: Mycoplasma hominis  
C:Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 11-Jan-2002  
C:Accession: PC6003  
R:Labeled: S.A.; Jensen, L.T.; Brock, B.; Birkeland, S.; Christiansen, G.  
J. Bacteriol. 178, 2775-2784, 1996  
A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis Imp gene system  
A:Reference number: JCB009; MUID:96213016  
A:Accession: PC6003  
A:Molecule type: DNA  
A:Residues: 1-624 <LAD>  
A:Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64859.1; PID:g1197337  
C:Genetics:  
A:Gene: Imp4  
A:Genetic code: SGC3  
C:Superfamily: surface-located membrane protein Imp3; tetratricopeptide repeat homology  
A:Keywords: duplication; membrane protein

	Query Match	5.6%	Score 133	DB 2	Length 624;	
	Best Local Similarity	22.4%	Pred. No. 0.39;			
	Matches	114;	Conservative	85;	Mismatches	171; Indels 138; Gaps 30.
Oy	2	APDSVGDVTKTLTLLASESDVSANAYMINSDMTLSAVSDNFAERICQYPKGSGCSA	61			
Db	98	AOADMLARSTKQL--NKSI--SSANTL--AKLTD-----KNTIQAAVTELEK-----	141			
Oy	62	SVSAYSRCACODCTLTLOSILKYPLEAKYOPPLTPDYOLEAFLTFKESDAMPANSTERK	121			
Db	142	EYOKANAQVAASINTNASMGSAKSISDAKYTEI-----KKLE-IFNKDKDAKFELQOTRKN	196			
Oy	122	FWMRRRGKNHSYFDLVFNLEK-----NTRDDATDIENFPASKRYLYMATLYKTYTN	176			
Db	197	IDFINTNKTNPYSELLISQTSKRDSKSNYSNNKSIDIER-AMTELQAOLA--KAND	253			
Oy	177	VDEFGASEFNKLSPFTTGJFGMGIRALKIKIIRSNLPIDIGTHSVSRLOHITSSYKDYM	236			
Db	254	KDQ-----ADNLTA-----KSKEDL-----NNSVSANTYLAKLTDRKN	287			
Oy	237	TQIPALPFAFKFSLMVOORLLATYAAGVDTFWYKKWIKLKLNEMVRNFVPTPKF--FN	294			
Db	288	TIOAQKTELEK---VOKANOVAASNTAS---MQSAKSSLDARKYTEIKTKLETEN	337			

```

Oy 295 -----KEIPEPSKALKE---KSTP-----TKLGFENKIGOGVDFPN---KEI 332
Db 338 KDKAKREKLEQTKKDIDFELKQIENDPQTKKNYQNIYKNLKKRKAKEKNSITFSSNNKKEI 397
Oy 333 RDPSPKALKEKVSNDAKDLFENKI-GOGTVDFINNEIRDPSPKALIRKVSFGAEDELFE--- 387
Db 398 QDANKSLQDELNN-----AKITKKGITDFVYSK-----KQL-----EDLKITDQA 437
Oy 368 NKIGOGTVDFI-----NNEIRDPSPKAL---IRKVV---TEADDLFENKIGQ----- 427
Db 438 KKVQTTTEADPTILDIHYKNISDASKNEELKQATQKINDIKIITEFKIOEKRRNEFSOFEQIK 497
Oy 428 -GTVDFTINKEIRDPSPKALIRKVSFEADN 454
Db 498 NELQSFINKDLKDOKVYSIR---TKIEN 522

```

RESULT 12  
S48326  
hypothetical protein YML071c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 02-Dec-1994 #sequence\_revision 09-Mar-1996 #text\_change 29-Oct-1999  
A:Accession: S48326; S48825  
R:Bowman, S.; Churcher, C.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: S48326  
A:Accession: S48326  
A:Molecule type: DNA  
A:Residues: 1-113 <BOW>  
A:Cross-references: EMBL:Z38114; NID:g558402; PIDN:CAAB6249.1; PID:g558403; MIPS:YML0  
R:Brown, D.; Bowman, S.  
submitted to the EMBL Data Library, October 1994  
A:Reference number: S48816  
A:Accession: S48825  
A:Molecule type: DNA  
A:Residues: 88-607 <BRO>  
A:Cross-references: EMBL:Z46373; NID:g587529; PIDN:CAAB6507.1; PID:g587539; MIPS:YML0  
C:Genetics:  
A:Map position: 13L

```

Query March 5.4%: Score 127; DB 2; Length 607;
Best local similarity 22.1%: Pred. No. 0.92;
Matches 90; Conservative 63; Mismatches 147; Indels 108; Gaps 21;

QY 108 KESDANPANSTKREPMWFRRGKNHSEPHDVLVFNLLKNTVTDADATDI-----EN 158
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 EDNDTGRIMTTESN---MLARKKKEDEPHKALSRLRNRISRKEDDKDIDRSDTLVYVEN 193
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 159 FAS-----RYLYMATLYYKT-----YTVNDFEGASFFPKLSFT 191
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 194 LDSITDLMELPFLARTCTIRTGHYOEAVMLYHTTSSLSRSPGSTIVDVCERKVLNLS-T 252
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 192 TGLFGW-----GIRKALKOIIIRSNLPIDIGTEHSY-----SRLOHITSSYKDY 234
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 253 TMLSGVLKLTSTNVSNSLKLTL-QYLSNIPPFOSKTKMSSLVSFLAMRYFTIDEIASY 311
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 235--MDTQIPALPKFAKFSL--MYVQRLATVAGYDTPYKTKMYKLNKFM--VNRVFLEPT 289
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 312 PLDVE-----SSNESLIEKMYKRIEVLREHV---YMSLNVLKSFVLDVTDLELPT 359
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 290 KKFENKEIREPSKALKEVSTDTKDLFENKIGQGVDFENKEIRDPSKALKEKYSNDKD 349
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 360 ---FEEELF--STVLRKINGTNEKEIEKE-----KETKKEEYQKODSVAANNED 404
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 350 LFEKNTGGQYDVFINNEIRDPSKALIRKVSYGAEDLFENKIGQGVDFINNEIRDPSKAL 409
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 405 YLENK-----SIEDVQEEV---QGVKEGEDDGAERKTEENIENETV---KKTEDKAE-- 450
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 410 IRKYVTEADDLFENKIGQYDVFINKEL--KDPKALLIRKYSTEDNL 455
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 451--KEKEEENKTRDKNAEKEEEEENKVEVTPPEPSKSDINKKKEEEL 496
      ::::: ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

```



OY 209 SNLP-----LDIGTE---HSVSRLOHTTSSYKDYMDQIPALPK-----FAKRRS 250  
DB 134 RSLPDMADGYALGIGIAEAKRIHQAVNNPGRSFRDNKRLISLDDGFDSFEKEHFT 193  
OY 251 LM--VVORLATVAGYVDPYWK--KW---YMKLNKPNVNRVFIPTKFKFKEIRPSK 302  
DB 194 FLQSYVMQDLTKLGYEYTTIDGQKIGCMGNGIINDLYKSVKREMYGIFEIYVNNIKQNE 253  
OY 303 ALKEKVS-----TDTKDLFENKIGQGVDFENKKEIRPSKA 338  
DB 254 AFKKEINSLVDMKAAGKEFDDLTQNNNLQAAEIIYNDVDTSGQIEKGV-AIKE 312  
OY 339 LKEKVSNDKULFENKIGQGVDFINNEIRPSKALIRKVSNGAEDLFENKIGQGVDFI 398  
DB 313 LSEKKKNAASDLADS-----AEKAKQVEDLQAQAKAYEN--AKSTAERA 357  
OY 399 NNEIRDPKAL--IRKYTTEADDLFENKIGQGVDF 431  
DB 358 AQAAREFFKGLPSFKDLAEKFRDLFPNP--EGMID 390

## RESULT 16

HAIYVUR  
hemagglutinin precursor - Influenza A virus (strain A/USSR/90/77)  
C:Species: Influenza A virus  
C:Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 16-Jul-1999  
C:Accession: A04064  
R:Concannon, P.; Cummings, I.W.; Salser, W.A.  
J. Virol. 49, 276-278, 1984  
A:Title: Nucleotide sequence of the Influenza virus A/USSR/90/77 hemagglutinin gene.  
A:Reference number: A04064; MUID:84090410  
A:Accession: A04064  
A:Molecule type: genomic RNA  
A:Residues: 1-566 <CON>  
A:Cross-references: GB:K01330; NID:9324193; PIDN:AAA43206.1; PID:9554652  
C:Genetics:  
A:Map position: segment 4  
C:Superfamily: Influenza virus hemagglutinin  
C:Keywords: hemagglutinin; homotrimer; 11proteins; thiolester bond  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-34/Product: hemagglutinin HA1 chain #status predicted <HA1>  
F:345-566/Product: hemagglutinin HA2 chain #status predicted <HA2>  
F:555,562/Binding site: palmitate (Cys) (covalent) #status predicted

## Query Match

Best Local Similarity 5.1%; Score 121; DB 1; Length 566;  
Matches 94; Conservative 68; Mismatches 161; Indels 130; Gaps 23;

OY 33 DMSDYLAVS--DNF---AERICQVPGKSNCSASVSAVMSRCAKOD-----CLTQOS 80  
DB 117 ELREDLSSVSSEFEREIPKKE--SWPK-HNVTGVTASCHKKSSFTYRLMLLWTEKN 172  
OY 81 LKYPLEAKYQPLTPDPYQLEAFLFKESDA-----NPANSTEKRFMRFRGKNHS 133  
DB 173 GSYF-----NLSSSYNNKEKEVLVLAVGHHPENIDQTYR---KENA 214  
OY 134 YFHDLVFLLEKNTYRQADADIENFASRYLMAFLYKTYTNVDFGASF---FNKLSF 190  
DB 215 YVSVAVSNNRFRPELIERPKRVNGOAGRINYWTLEPQDTIIFEANGNLAPNHAFL 274  
OY 191 TTGLGWMGIRKALKOIISNLPLODTEHSVSRLOHTSS--YKDYMDQIPALPKPAKR 248  
DB 275 NRG-FGSG-----ITISMSMDCDTRKCOPTGCAINSSLPFONIHPTVGTGECFKNYS 326  
OY 249 FSLMVV-----ORLATVAGYVDPW---YKKMY-----M 275  
DB 327 TKLRVAVTGLRNIPSTQSGGLFGALGIEGWTGIDGMWGYGHQNEGSGVADQKSNQ 386  
OY 276 KLNKPNVNRV-----FIPTKFKFKEIRPSKALKEKVSITDID-----L 315  
DB 387 NAINGITKKNVSVIEKMTQPTAVGKEFNK--LEKREMNLNKKVVDGDFLDIWTYNAELLVL 445

OY 316 FENKIGQGVDFENKKEIRPSKALKEKVSNDKDLFENKIGQGVDF---INNEIRPSK 372  
DB 446 LEN-----ERTLDPHDSVKNAYEKVSQLKNNAKE-----IGNGCFEFYHKCNE----- 491  
OY 373 ALIRKVSNGAEDL---FENKIGQGVDFINNE 401  
DB 492 -CMESVKNQTYDYPKYSSESKLNREKIDGVKLE 523

## RESULT 17

D71623  
erythrocyte membrane protein PfEMP3 PFB095C - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Jul-2000  
C:Accession: D71623  
R:Gardner, M.J.; Jettellin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600; MUID:99021743  
A:Accession: D71623  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2441 <GAR>  
A:Cross-references: GB:AE001371; GB:AE001362; NID:93845092; PIDN:AAC71809.1; PID:9384  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB095C

## Query Match

Best Local Similarity 5.0%; Score 118.5; DB 2; Length 2441;  
Matches 58; Conservative 32; Mismatches 64; Indels 31; Gaps 14;

OY 294 NKEIR-EPKSKALKEKVSITDIDFENKIGQ---YDFENKKEIRDP-SKALKEKVSNDK 348  
DB 1121 NKELOKNGSEGLKENAELKNKEL--QNKGSGLKENAELKNKELRNKSGLEKENAELKNK 1179  
OY 349 DLFENKIGQGVDFINNEIRDPKALIRKVSNGAEDLFE-----NKIGQ---TVDFI 398  
DB 1180 EL-QNKGSGLKE--NAELK--NKELOKNGSEGLKENAELKNKELONKSGLEKENAELK 1234  
OY 399 NNEIRDP-SKALIRKYTTEADDLFENKIGQGVDFINKEIRPSKALIRKVSSTEAD--- 453  
DB 1235 NKELRNKSGLEKENYVTN--NDLKNNDI--QNKDLSNKKMK--NKELLNKDISNKKMK 1289  
OY 454 NLEK 458  
DB 1290 ELLNK 1294

## RESULT 18

A72358  
conserved hypothetical protein - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: A72358

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
Garrett, M.W.; Stewart, A.W.; Colton, M.D.; Pratl, M.S.; Phillips, C.A.; Richardson,  
C.M.  
Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MUID:99287316

A:Accession: A72358  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-447 <ARN>

A:Cross-references: GB:AE001734; GB:AE000512; NID:94981107; PIDN:AA035685.1; PID:9498  
A:Experimental source: strain MSB8

C:Genetics:  
A:Gene: TM0600

Query Match 5.0%; Score 117.5; DB 2; Length 447;  
Best Local Similarity 21.8%; Pred. No. 2.6;

Matches 74; Conservative 50; Mismatches 104; Indels 111; Gaps 20;

OY 120 KRFRMR-FRRCKNHSYFHDVLENLEKVNTRDADATDIENFAS----RYLYMATLY----- 170  
122 KRFRHNGEPYKSY-FRH-----LCGDILKETKKNITYEFSSENEFFLGCTVEDVEF 174  
OY 171 --YKTYTNVDEFGASFENKLSFTTGLFGWCIGIKRALK---QIIRSNLPDIDGTEHSYSR 223  
175 GPRHVRHTNRDDE---FDELIVAVGRSGHRLMERLEKEKYPQLVRPNQFVIDIGRYELP- 230  
OY 224 LQHTSSYKDMQOI-----PA-----LPFAKRFSLMAYVQRLATYAG 263  
231 -NHMDPFSMDYEVKVRFRKTYGICRFGCPNPAKVTLEKED-----FTTVNG 279  
OY 264 YVDTPWKYKMYMKLNEMVNRVFIPTKKFENKEIREPS-----KALKEKYST 310  
280 YSDS-----LHKTEN--TNPAILVTR--FTPEPKDPGIVGNLAKLANILAGDKREKYL 330  
OY 311 DTKDLFENKIGQGTVDFFENKEIR-----DPSKALKEKYSNDADKDLFENKIGQTVDF 362  
331 QTVG-----DF--KEFRRTKRLGRVHPFLDSESYILGDANLVPSPKIRESLVDF 377  
OY 363 INNEIRDPKALIRKVGTCGAEDLFENKIGQTVDFINNE 401  
378 VEN-----LEKVIPLGA-SYFDNLLYAVEVKEFTYTK 406

## RESULT 19

H90565  
Restriction modification enzyme subunit r2 [imported] - Mycoplasma pulmonis (strain UAB C/Species: Mycoplasma pulmonis  
C/Date: 24-May-2001 #sequence: revision 24-May-2001 #text\_change 03-Aug-2001  
C/Accession: H90565  
R/Chamaud, I.; Heilig, R.; Ferris, S.; Barde, V.; Samson, D.; Gallsson, F.; Moszer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A/Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm  
A/Reference number: A9512; MUID:21261165; PMID:11353084  
A/Accession: H90565  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-986 <KUR>  
A/Cross-references: GB:AL445566; PID:G14089846; PIDN:CA013605.1; GSPDB:GN00153  
A/Experimental source: strain UAB CTIP  
C/Genetics:  
A/Gene: MIPU\_4320  
A/Genetic code: SGC3

Query Match 5.0%; Score 117; DB 2; Length 986;

Best Local Similarity 23.1%; Pred. No. 7.7; Mismatches 142; Indels 60; Gaps 17;  
Matches 81; Conservative 67;

OY 140 FNLEKVNTRDADATDIENFASRYLYMATLYKTYTNVDEFGASFFN--KLSFTTGLFGW 197  
648 FSLREGGIN--DAFKIYANSSDKRIQOLVYGEQVDEDFINMWSLKISFS----- 698  
OY 198 GIKALKQIIRSNLPDIDGTEH--SVSRLOHTTSSYK-----DYMDOIPLALP 243  
699 NIDDEKNEEFIRNLSLENKKYILNLSQVSNIFSSLTKFEYGENEKISDSLSQLOMNOY 758  
OY 244 KFA-----KRFSLAYVORLATVAGVDTPMYKMYMKLNEMVNRVFIPTKKFENKEIR 298  
759 KYAIEIKKLNSTNEKEKISYVLNSIDISNIFAY---KEKIIDETIYENLLEFNKKISK 815  
OY 299 EPSKAL--KEKYSTDTK--DLFENKIGQGTVDFFENKEIRDPKALKEKYSNDADKDLFENK 354  
816 YPNRLTYEDTLSEIDKHILQIKNNYNGKI--NOKYEYFLLVQWKKEIKNFIFIKK 872  
OY 355 IGGQTVFINNEIRDPKALIR-----KYSTGAEDLFENKIGQGTVDFFINNEIRDPKALI 410  
873 --DKSLD--EKFTIDYKRLIKSVQKVKNOIEAMWLEKIKVEYHGGINDDIRDMKKRI 928  
OY 411 RKVYTEADDLFENKIGQGTVDFINK---EIRDPKALIRKYSTEADNLE 457

Db 929 ND--KDLDIDKS-----EFIKWMSRKSREKVDKIIDKLSIEKESIE 969

## RESULT 20

B42641  
Kinesin-related protein C1N8 - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein YEL061C  
C/Species: Saccharomyces cerevisiae  
C/Date: 03-Mar-1994 #sequence: revision 19-May-1995 #text\_change 23-Mar-2001  
C/Accession: S50528; B42641; S44939; S69009; S20134  
R/Dietrich, F.S.  
Submitted to the EMBL Data Library, December 1994

A/Description: The sequence of S. cerevisiae cosmid 9669, 8334, 8199, and lambda c10

A/Reference number: S50434  
A/Accession: S50528

A/Molecule type: DNA

A/Residues: 1-1038 <DIE>  
A/Cross-references: EMBL:U08795; NID:9603241; PIDN:AAB65026.1; PID:9603257; GSPDB:GNO

R/Hoyt, M.A.; He, L.; Loo, K.K.; Saunders, W.S.  
J. Cell Biol. 118, 109-120, 1992

A/Title: Two Saccharomyces cerevisiae kinesin-related gene products required for mito  
A/Reference number: A42641; MUID:92317149  
A/Accession: B42641

A/Molecule type: DNA

A/Residues: 1-253, 'A', 235-830, 'H', 832-1038 <HOY>

A/Cross-references: EMBL:Z11859; NID:93541; PIDN:CA077885.1; PID:93542  
A/Note: sequence extracted from NCBI backbone (NCBIIN:107723, NCBIIP:107726)

R/Roussellet, G.; Simon, M.; Ripoch, P.; Buhler, J.M.  
Submitted to the EMBL Data Library, May 1994

A/Description: A second nitrogen permease regulator in Saccharomyces cerevisiae.  
A/Reference number: S44938  
A/Accession: S44939

A/Molecule type: DNA

A/Residues: 1021-1038 <ROU>

A/Cross-references: EMBL:X79105; NID:9485969; PIDN:CA055722.1; PID:9485971  
R/Roussellet, G.; Simon, M.; Ripoch, P.; Buhler, J.M.  
FEBS Lett. 359, 215-219, 1995

A/Title: A second nitrogen permease regulator in Saccharomyces cerevisiae.  
A/Reference number: S69008; MUID:95172238

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1021-1038 <ROU>  
A/Cross-references: EMBL:X79105; NID:9485969; PIDN:CA055722.1; PID:9485971

A/Genetics:

A/Gene: SGD:CIN8; MIPS:YEL061C  
A/Cross-references: SGD:S0000787; MIPS:YEL061C

A/Map position: 5L

C/Superfamily: kinesin-related protein C1N8; kinesin motor domain homology  
C/Keywords: ATP; coiled coil; microtubule binding; mitosis; nucleotide binding; P-loop  
F/16-173/Region: nucleotide-binding motif A (P-loop)  
F/112/Binding site: ATP (Lys) #status predicted

Query Match 5.0%; Score 117; DB 1; Length 1038;

Best Local Similarity 19.5%; Pred. No. 8.2; Mismatches 214; Indels 180; Gaps 28;  
Matches 119; Conservative 97;

OY 4 SDSGVDTKTLAASEVSDAANAYMI-----NSDMSYLSAVSD-----NFA 46  
188 SDAGIIRVLKLPDLELOQNDYVVCSTFELYNEKLKLDLSNSGSSNTGFDGFM 247  
OY 47 ERICQVPPKGSNCASASVAYWSCAKODC-----LTLOS--LKYPLEAKYQPLTLPDP 97  
248 KKLRFDSSTANNTTNSGSSRSNSRNSPRSLNDLTPKALLRKLRTKSLPNTIQO 307  
OY 98 YOLEAFLFRESDANPANSFE-----KRFMRFRGKNHSYFHLY- 139  
308 YQQQAAVNSRNSSNSGSSNTNNASSNTNTNNGRSSMAPDQNGITYIONLOEPHTNA 367  
OY 140 --FNLEKVNTRDADATDIENFASR--LYMATLYKTY-----TN 176  
368 MEGNLLOKGLKHQVASTKMNDFSSRSHTIFTITLVKKHODELFRIKMNVLVDLAGSEN 427



```
QY 177 VDEFG-----ASFENKLSFTTG-----LFGWGIKRA 202
DB 428 IIRSGALNORAKEAGSINSGLTGLGVINALVDKSGHIFPRESKILRLLODSDIGGATKRA 487
QY 203 LKQIIRSNLPDIDIGHEHSVSRLOHTTSKYDMDTOIPALPKPAKRESLM-----V 253
DB 488 L---IATISPAKVTSEETCSTLEY-ASKAKNIKNK--POLGSEFIMDIIVKKNITMELAKI 541
QY 254 VQRLATATAG---YVDTPMYKKMYKMLKMFVNRYFIPKKKFNKIREPSKALKKVKST 310
DB 542 KSDLSTSKSEGIYMSQDHYKNLNSDLESYK-NEV-----QCKREIESLSKKNAL 591
QY 311 DTKDLFENKIGGTVDFPKKEIR-----DPSKALKEKYSNDAKDL--FENKIGOGT--- 359
DB 592 LYKDKLKS---ETIOSQOCQIESLKTTHDLRAQDKQKHKEIETSDPNKKLOKITEWA 648
QY 360 -----VDF-----INNEIRDPKALIRKYSTGAED--LEFNKIGOGTVDFIN 399
DB 649 QMALHDYKRELDLNOKFEMHITKEIKLSTFLQNLNMOQESILQETNI--QPNLDMIK 707
QY 400 NEIRDPKAL-----IRKYVTEADDLFENKIGGTVDFINKEIRDPKAL---I 445
DB 708 NEVTLMTMQKAEIMYKDYCKVKKILNESPKPEFNVVIEK--IDILRVDFQKFFYKKNIAENL 765
QY 446 RKVSTEADNL 455
DB 766 SDISEENNMM 775
RESULT 21
E81086
Iron-regulated protein FRPC NMB1415 [Imported] - Neisseria meningitidis (strain MC58 ser
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 Sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: E81086
R:Jettelajn, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamatheyan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizsa, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755
A:Accession: E81086
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1829 <TEXT>
A:Cross-references: GB:AE002490; GB:AE002098; NID:g7226651; PIDN:AAFA1776.1; PID:g722665
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1415
Query Match 5.0%; Score 117; DB 2; Length 1829;
Best Local Similarity 20.2%; Pred. No. 17;
Matches 81; Conservative 62; Mismatches 150; Indels 108; Gaps 19;
QY 94 LDDPQLEAFLTFKESDANPANSTFKRFWMFRGKNHSYHDLVFNLEKNVYTDADA 153
DB 35 LDDRRADDAALVIGKDAANGLN---LWM--KKG-----VENLMDVVGK--K 75
QY 154 TDIEFNASRYLMAATLYKTYT-----NDEFGASFENKLSFTTGLEGWCIKRA 202
DB 76 TLLEKFDR---VALOHFOYARLINONNGRLPNTSEIERSYKAVTDN---GVSSAA 127
QY 203 LKQIIRSNLP-----LDIGTE---HSVSRLOHTTSKYDMDTOIPALPK----- 244
DB 128 IDLVINRSLPDMADGYWALGLIGIEARINHEQAVNPNPNCESERNKQILSALDKGDFGSF 187
QY 245 FKKRSLMAYVQQLLATVAG--YVDTPMYK--KW---YMKLNFMVNRVFIPIKKFFNKE 296
DB 188 KKHHTFLOSVMDVYKGLVEYTTIDGMOKIGGNGCIINDLKYSVKRMTGTGFEIIVNNN 247
QY 297 IREPSKALKEKYS-----TDTKDLFENKIGGTVDFENKEI 332
```

```
DB 248 IKOGNEAFKNEIISLVHDKMAAGKEFGDDLNTQWNNNLTOAAEIIYNDIYDNTSOGIEKGV 307
QY 333 RDSKALKEKYSNDADLFEFNKIGOGTVDFINNEIRDPKALIRKYSTGAEDLFEFNKIQ 392
DB 308 K-AIKELSEKMKMAASDLADGS-----AEKKQVEDIAQAKAEYEN--AK 351
QY 393 GTVDFINNEIRDPKAL--IRKYVTEADDLFENKIGGTVDF 431
DB 352 STAEKAAQAAREFFKGLPSFKDLAEKFRDLFRP--EGWID 390
RESULT 22
G70163
hypothetical protein BB0512 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: G70163
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: G70163
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2166 <KLE>
A:Cross-references: GB:AE001153; GB:AE000783; NID:g2688419; PIDN:AAC66876.1; PID:g268
A:Experimental source: strain B31
Query Match 5.0%; Score 117; DB 2; Length 2166;
Best Local Similarity 19.3%; Pred. No. 21;
Matches 105; Conservative 94; Mismatches 182; Indels 164; Gaps 24;
QY 16 AAESVDSANAMVIMNSDSDYLSAASVDFAEIRICGVFKGSCASVSAYMSRCAKQDC 75
DB 500 ALFESIDSSSKFE--NOMESKYKS-----FTDKLTAGMDEFSLMVGKEPTLSQEAATNNY 553
QY 76 LTLQSLKYLEAKYQPLTPDPQLEAFLTFKESDANPANSTFKRFWMFRF----- 128
DB 554 QERQDNLKLE-----NEIESFYMFKE-----TOETLKVDFNTSLINIKD 594
QY 129 --GKNHSYFHDVFNLEKNVNR-----DADATDIENFASRYIYATLY 171
DB 595 ETGKNVIEFRDRYDEVNIFVTOLESKLOYSKMOGEMSNLKNISSQINK----- 645
QY 172 KTYTND-----EFGASFENKLS-----FTTGLGF--WGIKRALQIIRSNLP 212
DB 646 ---TNEEFLSLIOIKDKGIELSESVFNDSIHTQKKAIDMHGSW--KDELIALKSLLD 700
QY 213 LDIGTEHSVSRLO--HTTSKYDMDTOIPALPKFAKRFSLMAYVQRLATVAGYVDTPMYK 271
DB 701 IKVSEELLSATLKTIESLEKDVNDMEYVLTKTGDIESLVIEK----- 744
QY 272 KWTMKLNFMVNR--VFIPTRKFFNKE---IREPSKALKE---KYSTDTKDLFENKIQ 322
DB 745 --YKELKDSYSSQDEAIIIGIKFEINRQTEIINDKSVFMLEDLNKKFDDKNNVISKIEE 802
QY 323 GTVDFENKIREPSKALKEKYSNDADLFEFNK-----GOGTV----- 360
DB 803 --CDYKLOFKISEEDILNNFKSDLNEFIESKQIYISNKSQKQIDDFLDISKDILN 860
QY 361 --DFINNEIRDPs--KALIRKYSTGAEDLFEFNKIGGTVDFINNEIRDPKAL----- 409
DB 861 RKDSINNEVDSKLSMDQSKLNETITVKTENLSS--GKVDLDLDSVYTTKIKELKFSIES 918
QY 410 IRKYVTEADDLFEN-----KIGOGTVDFINKEIIRDPKALIRKYST-----EA 452
DB 919 LESYLELEKIDEFNQAIAYSDELLDIDIMHFNKRETRLEENLSKKFPAVALNNSSEEFVEY 978
QY 453 DNILE 457
```

Db 979 DSLQ 983

RESULT 23

G97085

Sensory transduction histidine kinase (with HAM domain) [Imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001

C:Accession: G97085

R:Noiling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Dally, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: G97085

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-473 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK79474.1; PID:915024454; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1507

Query Match 4.9%; Score 116.5; DB 2; Length 473;

Best Local Similarity 21.8%; Pred. No. 3.2;

Matches 86; Conservative 62; Mismatches 120; Indels 127; Gaps 20;

Db 86 EAKYQPLTPDYOLEAFLEKESDANPANKTERF-----WMRFRGKNHSTFH 136

16 EAFYVPSI---YLINSFKSLNLODEISSGISEKRFSSLIQANLTFKIRKRGK-----Y 67

QY 137 DLVFNLEKNVTRDADA---TDIENFASRYLYMATLYKYTYNDEFGASEFNKLSFTTG 193

Db 68 DVSXVLTKE---ELDVFINTYLSDFKQKYIDV-----IDEKNSVFNDS----- 111

QY 194 LFGMCIRKALQIIRSNPLDIGTEHSVSRQHTSSYKIDMDQIPALPKFAKFSIMV 253

Db 112 -----NKLPRKREPLDNKNQKYLINDING--KTY-----LFI 143

QY 254 VQRLATVAGVDPFPMYKKMYMK-----LKNFMVNRVFIPTKKFNKE 296

Db 144 VSKL-----KLDNSYKFSYIKDYSKIYDNKRYLLNLLKINRITIIILIIIMILSKF 197

QY 297 IREPSKALKEKVSFTDKDLFENKIGQGVDFPNKEIRDPKALKKVSNDADKLEFNKIG 356

Db 198 IYKPLNVMIKST-----QKIEGN--FN-----ERVSGVRND---EIG 230

QY 357 OGTVDFINNEIRDPKALIRKVSQAEGLFENKIGQGVDFINNEIRDPKALIRKYTE 416

Db 231 QLSKNF--NYMADVLEDKIKELKTSSED-----KQRPIDDLTHEIRTPLSII---GY 278

QY 417 ADLFEKNIGQGVDFINKEIRDPKALIRKYSTE 451

Db 279 ADPLRTAKYDEKTLFSSINVIYDECKRL-QKLSSK 312

RESULT 24

G82925

hypothetical protein UUI62 [Imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: G82925

R:Glass, J.I.; Letkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

A:Title: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit

A:Reference number: A82870

A:Accession: G82925

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-457 <GLA>

A:Cross-references: GB:AE002116; GB:AF222894; NID:96899118; PIDN:AAK30568.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: UUI62

A:Genetic code: SGC3

Query Match 4.9%; Score 116; DB 2; Length 457;

Best Local Similarity 19.4%; Pred. No. 3.3;

Matches 83; Conservative 61; Mismatches 133; Indels 150; Gaps 19;

QY 78 LOSLKYPLEAKYQPLTPDYOLEAFLEKES---DANPNANKTERFWMRFRGKNHSTY 134

Db 25 LEDNQLERKKQALT-----TFILRAQSMIDCVASNG-----GAGFYISW 66

QY 135 FHDVFNLEKNVTRDADAT-----YNNVDEFGASEFNKLSFTTGLGWMGKRAK 204

Db 67 FDDLKTKTDYQKTKFEALATLNYLWLOGHFIDKDYSAIGDSLSEKSENADPEKRL 126

QY 155 DIENFASRYLYMATLYXT-----YNNVDEFGASEFNKLSFTTGLGWMGKRAK 204

Db 127 DIIN---KLKRCGLYKSKFANFNDDGIYVYDELKINAIN-----NVHEELR 171

QY 205 Q-----IIRSNPLDIGTEHSVSRQHTSS-----YRDYM--DQIP--ALPKFA 246

Db 172 DHYISKNNPVVNNFDLN---HHNINVSINASGNLATLNGFILNDOCTLPRLDITDKSN 228

QY 247 KRESLMVYQRLATVAGVDPFMYKK--WYMKLKNVNRVFIPTKKEIRDPKAL 304

Db 229 KYVD-----NATANYVDIHDVKNLAWYGOONNNAITR-----DELYEYCNKI 270

QY 305 KEKVSFTDKDLFENKIGQGVDFENK-----EIRDPKALKKVSND 345

Db 271 NSKMTWDA---IYDKAGELYVDLDEKTIYKYKALVNNQSGTHAIYVPSKQINENVEL 328

QY 346 -DAKDLEFNKIGQGVDFINNEIRDPKALIRKVSQAEGLFENKIGQGVDFINNEIR 404

Db 329 INIQIDINKIKEGVQADVKNVNEVKAIPTKEVTFKNL-ENKVNBAKQAVLSSYBOD 387

QY 405 PSKALIR 411

Db 388 KASGVFK 394

RESULT 25

E81339

probable restriction/modification enzyme Cj0690c [Imported] - Campylobacter jejuni (

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 31-Mar-2000

C:Accession: E81339

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chl

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals

A:Reference number: A81250; MUID:20150912

A:Accession: E81339

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1250 <PAR>

A:Cross-references: GB:AL139076; GB:AL111168; NID:96966128; PIDN:CAE72964.1; PID:9696

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj0690c

Query Match 4.9%; Score 116; DB 2; Length 1250;

Best Local Similarity 19.9%; Pred. No. 12;

Matches 81; Conservative 56; Mismatches 145; Indels 126; Gaps 15;

QY 134 YFHDVFNLEKNVTRDADATDIENFASRYLYMATLYKYTYNDEFGASEFNK 185

Db 430 YFDNDFKILK-----SKVKEYT--FYKKGOIYKNTNSRKASASFTYPOSTANF 480

QY 186 -----NKLSTP-----TGLFGMGIRKALQIIRSNLP-----LDIGTEHS 220

Db 481 LIQSALRKLNNEIRKLIIDNACGSGHFLVGVAINATHIVLSDFHTNLKELYEERK 540

OY 221 VSRLQHTTSSYKDYMDQIPALPKFAKRESLMVGORLATVAGYVDPWYKWKMLKNE 280  
DB 541 EMLHNIKOFVODYEDESDILKRL-----LAKRIYGVOLNPSIELTSLM 589  
OY 281 MVRNVPITPKKFNKEIRPSKALKKESVTDTKDLFENKIGQTVDFENKEIRDPKALK 340  
DB 590 IDSFJFTPLSFLEHIKICGNALINSNLS-DEKDL-----IKONSSMLFNSTIQEPEILQ 644  
OY 341 -----EKVSN-----DAKDLFENKIGQ-----GTVDFINNEIRDPKALK 375  
DB 645 EYFELKLDNKDITNEQIKOSKOIYQNETPKLNKLMLYNTLTIFHVKEELQILKAL- 703  
OY 376 RKVSTGAEDLFENKIGQTVDFINNEIRDPKALKIRKVTYDAD-----DLFENKIG 426  
DB 704 -----SQD-----DLONLSQNOQAALISKYQKEFFEFYEELEFPFIVEVNOVF 746  
OY 427 QG-----TVDFINKEIRDPKALKIRKVTSTADNILEK 458  
DB 747 KGFIIIGNPMWDKTKFSDSDFEPQYKSDYRSLIASKKEIDONILAK 794

## RESULT 26

hypothetical protein jhp0440 - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 28-Jul-2000  
C:Accession: H71931  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merber, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A:Reference number: A71800; MUID:99120557  
A:Accession: H71931  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-912 <ARN>  
A:Cross-references: GB:AE001478; GB:AE001439; NID:g4154971; PIDN:AAD06018.1; PID:g415497  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: jhp0440  
C:Superfamily: Helicobacter pylori hypothetical protein jhp0440

Query Match 4.9%; Score 115; DB 2; Length 912;  
Best Local Similarity 20.5%; Pred. No. 9.4;  
Matches 105; Conservative 51; Mismatches 167; Indels 188; Gaps 24;

OY 98 YOLEAFAFLFKESDANPA---NSTEKRFWMRFRRGKNHSYFHDVFNLEKNYTRADAT 154  
DB 312 YOLDVAIV--KAKDLKPSFTGTGOTKRTDM-----NEQIKSIANFPDKKIFSGGFE 363  
OY 155 DIE-----NFA--SRILYATL--YKTYTANDE-----FGA 182  
DB 364 DLPILHDCQVLAGNHRIGAMLNFPKSRYYNKAIKEYYRIDLKDDELVRVPHORLDN 423  
OY 183 SPFNKL--SFTTGLFCWGKIKRA-----LKOIINSNLPID 215  
DB 424 TELNNLAASSNGRPNSESDHAIVLSHYEAKIKELDKLDADISTYSLKITYAKKNLNFDR 483  
OY 216 GTEHSVSLQHTTSSYKDYMDQIPALPKFAKRESLMVGORLATVAGYVDPWYKWKMY 275  
DB 484 ATHPNVG-----DSNLALMYNMPRTKTQGIELL-----NMQKEFSN 521  
OY 276 KAKNF--MVRNVPITPKKFNKEIRF--PSKALK-----EKVSTDTKD 314  
DB 522 DIKSEYKVKMKEVDNAGSFHNLIDHNPVNSLNAVLSDIMRSPANLKNYPTSTSLSK 581  
OY 315 LFNENKIGQTVDFPNK-----EIRDPKALKKESVND--AKDLFE 352  
DB 582 LSEKFKTSSLDMEKSDSASDISILGATARFARFDDPSKALFESLSKSDIKKGLKD 641  
OY 353 NKIGQGVDFIN--NEIRD-----PSKA 373

DB 642 YK1ADTDKMFNDPSKEFKDIDIDYFTHYLLMVRREPNENNPNLRLIQAQKQKSESEK 701  
OY 374 LIRKVTGAEDLFENKIGQTVDF--FINNEIRDPKALKIR--KYVTEAD--LF--ENKIG 426  
DB 702 GIKVSKQAAEETEKIGIKRPSDTNLSNNEI---KALLNNAKIPISGRDAVIFGKNLN 757  
OY 427 QGTVDFINKEIRDPKALKIRKVTSTADNILE 457  
DB 758 PEIVERFIHK---NHKRAIFKASNKETIELQ 785

## RESULT 27

S49394  
HsdRI protein - Mycoplasma pulmonis  
C:Species: Mycoplasma pulmonis  
C:Date: 29-Nov-1995 #sequence\_revision 01-Mar-1996 #text\_change 07-Dec-1999  
C:Accession: S49394  
R:Dybvig, K.; Yu, H.  
Mol. Microbiol. 12, 547-560, 1994  
A:Title: Regulation of a restriction and modification system via DNA inversion in Myc  
A:Reference number: S49391; MUID:95020589  
A:Accession: S49394  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-986 <DYB>  
A:Cross-references: GB:L25415; NID:g496153; PIDN:AAA65632.1; PID:g496157  
C:Genetics:  
A:Genetic code: SGC3

Query Match 4.9%; Score 115; DB 2; Length 986;  
Best Local Similarity 23.1%; Pred. No. 10;  
Matches 81; Conservative 67; Mismatches 142; Indels 60; Gaps 17;

OY 140 FNLEKNYTRADATDIENFASRLYMATLYTYTNAVDFGASFEN--KLSTTGLFGW 197  
DB 648 FSLSEQOSIN--DAFIYANSSDKELIQOLYGCYGEQVEDFINFWSLKISFS----- 698  
OY 198 GIKRALKOIIRSNLPIDIGTEH--SVSRLOHTITSSVK-----DYMDQIPALP 243  
DB 699 NIYDEKNNELFRNISLENNKKYLLKNSOVNSIFSLTFEYKNEKISIPSLBLOMQO 758  
OY 244 KFA-----KRFSLMVGORLATVAGYVDPWYKWKMYKMFNVRNVPITPKKFNKEIR 298  
DB 759 KMANEIKKNSTNEKEKISTEVLNSIDISNIKRAY---KEMIIDEITLLENLFFPNKIKSK 815  
OY 299 EPSKAL--KEKVSVDTK--DLFENKIGQTVDFENKEIRDPKALKKESVNDADKLEFNK 354  
DB 816 YPNNRRLTYEDTLSEIDKHIOILKNNYNOGKI---NOKYEIIFLLVQKMKNEIKNFPIKK 872  
OY 355 IGGQTVDFINNEIRDPKALKIR---KVSIGARDLFENKIGQGVDFINNEIRDPKALK 410  
DB 873 --DKSLD--EKEFTIDGKRLKSVFQKVKNOIEAMWEKTLKEYHGGINNOQIRKDWKKRI 928  
OY 411 RKVYTEADLFENKIGQTVDFINK--EIRDPKALKIRKVTSTADNILE 457  
DB 929 ND--KDLDDIEKS-----EFIKKMSRSKKEVDKDIIDKLSIEYKRSIE 969

## RESULT 28

S38173  
myosin-like protein MLP1 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YKR095w; protein YKR415  
C:Species: Saccharomyces cerevisiae  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 21-Jul-2000  
C:Accession: S38173; S40647; S31207  
R:Baladron, V.; Ballesta, J.P.G.; Bou, G.; del Rey, F.; Esteban, P.F.; Garcia-Cantale  
submitted to the Protein Sequence Database, March 1994  
A:Reference number: S38158  
A:Accession: S38173  
A:Molecule type: DNA  
A:Residues: 1-1875 <BAL>  
A:Cross-references: EMBL:Z28320; NID:g486586; PID:g486587; MIPS:YKR095w  
A:Experimental source: strain S288C





Db 416 -----KQPLATQKDLTGVSTPVRQ--VKLQRADSRSLPAPKPLSSKSVRKE 464  
Qy 301 SK-----ALKEK--VSTDTKDLFENKIGQTVDFNKEIRDPSCA-L 339  
Db 465 SKKEAPEATKASQVEKTPVESKEKIVKKDKKGKVESKPSYTEKEVSPKAEV 524  
Qy 340 KEKVSNDADLEFNKIGQTVDFINNEIR-----DPSKALIRKVGTAEDLFENKI 390  
Db 525 AEKAATESKRP-----KVTKDKV--VKKEIKTPEEKKEPKKPKKAKEDKTPDKDEKPK 578  
Qy 391 GGGTVDFINNEIRDPSCALIRKYTTEADLFENK--IGQTVDFINKEIRDPSCALIRKY 448  
Db 579 KEBAKKEIKKEIKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEK 637  
Qy 449 STE 451  
Db 638 SKD 640

## RESULT 33

AB5176

hypothetical protein AT4g15890 (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001

C:Accession: AB5176

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: AB5001; MUID:20083488

A:Accession: AB5176

A:Status: preliminary

A:Residues: 1-1314 &lt;STO&gt;

A:Cross-references: GB:NC\_001268; NID:g7268337; PIDN:CA878631.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g15890

A:Map position: 4

Query Match 4.8%; Score 112.5; DB 2; Length 1314;  
Best local similarity 20.2%; Pred. No. 22;

Matches 99; Conservative 89; Mismatches 190; Indels 113; Gaps 27;

Qy 14 LLAASSVDSANAYMINDMSDYLAVSDNFAERI-----C-----SQVP 54  
Db 10 ILAGEGCGDESDYHVLVDLKSILDTDDDEILNRFYGLSSMASSFLRCISAMSPVE 69  
Qy 55 KGSNCASVSAYMSRCADKODCLTQSLKYPLEAKYQPLTLPDYOLEAPFILFKESDAMP 114  
Db 70 SGRALALADATLSTLLSTNC-----PVFTFSPIALFSLGSIIRRYLKRKRDSDAGO 121  
Qy 115 ANSTERFMMRRFRGRKNHSHYFDLVFNLEKNV--TRDADATDIENFASRYLWATLYYKT 173  
Db 122 GNSQGEKGNKKKRGK-----KRNLYGEDGETEGCGDAMLFIY----- 163  
Qy 174 YINVDIEFGA--SFENKLTSTGLFGWGIKRALKQIIRNMLPLDIGTEHS-----VSLQHI 227  
Db 164 ---LEKIGSVLSEFVHLDRPDSL-----KSLVQTV--SEIPL--LALEHSGVLNVDRLMEM 212  
Qy 228 TSSY-----KQYMDQIAPL-----PKFARFSLMVVQRLATVAGVVDNP 268  
Db 213 CGKILGGLVNSDHGMALTAELSKSLTPLLGMKIQAHSAFGLGFYSRKLMSLA--KDNP 270  
Qy 269 WYKKWYMKLNFMVNRVPIPTKK---FPNKEIRPSKALKEKVSVDTFDLFENKIGQTV 324  
Db 271 ELKKVYVSNLPKFLVHKA--PEKAEPGRGFAVEAVLEIVKAMEVGEQSEFVD--FVKKKCGCK 327  
Qy 325 VDFPKKEIRDPSCALKEKVSNDADL-----FENKIGQTVDFINNEIRDPSCALIRKVS 379  
Db 328 SNFRVLAV--DIPLLISSIGNPLGDISSENGLKGISGLICIDALVORCSDTS--ALIR--- 382  
Qy 380 TGAEDLFEKKGIGQTVDFINNEIRDPSCALIRKYTTEADLFENKIGQTV--DFINKEIR 438

Db 383 --ARAL--SNLAQ--VVEFLSGDER--SRSLKQALGPNGETSE---GKAVYDLDLKKRCV 432  
Qy 439 DPSKALIRKVS 449  
Db 433 D-EKAAYRRRA 442

## RESULT 34

T38077

hypothetical colled-coil protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T38077

R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V. submitted to the EMBL Data Library, April 1996

A:Reference number: 221767

A:Accession: T38077

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1957 &lt;CON&gt;

A:Cross-references: EMBL:Z70690; PIDN:CAA94624.1; GSPDB:GN00066; SPDB:SPAC1F3.06C

A:Experimental source: strain 972h-; cosmid c1F3

C:Genetics:

A:Gene: SPDB:SPAC1F3.06C

A:Map position: 1

Query Match 4.7%; Score 112; DB 2; Length 1957;  
Best local similarity 19.2%; Pred. No. 39;

Matches 107; Conservative 94; Mismatches 204; Indels 152; Gaps 27;

Qy 9 DVTKTLAASEVDSANAY--MINDMSDYLAVSDNFA-----ERISQVPSKSN 58  
Db 205 DISRQLLTVEKIDKKEKEDYKIKEDVSSIKASIAEQASNNKSLRGEORL--EKLIVSSN 263  
Qy 59 CSASVSAYMSRCADKODCLTQ-----SLKYPLEAKYQPLTLPDYOLE 101  
Db 264 KTVSTLRQTENSLRACCKTLQEKLEKCAINEDSKLLEELKHNV--ANYSDATVHKKLTIE 322  
Qy 102 AAFILFKESDANPANSTERFMMRRFRGR-----NHSYFHDVFNLE 144  
Db 323 DISTRISD-----NKSEEDLTSLNKEKLEKLRNTIGSLKSDSRNSQLEBEVEKLE 378  
Qy 145 KAVTRADATDIENFASRYLWATLYYKTYTNVDGASFPKLTSTGLFGWGIKRALK 204  
Db 379 SNRTIHSOLTDEASKLSF---EQENKSLKSIDEQ---NLSKSKDM---VQVSS 427  
Qy 205 QI--IRSNLPLDIGTEHSYSLQHTSSYKVDWDTQIPLPFAK--RFSLMVYQRLAT 260  
Db 428 QLEEARSSL-----AHATGKLAELNSE--RDFONKRIKPEKIEQDLRACLNSSELKE 480  
Qy 261 VAGYVDTPYKKWYMKLNFMVNRVPIPTKKEFNKEIRPSKALKEKVSVDTR--DLFEN 318  
Db 481 KSLALID-----KNDQELNML--REQIKQKKVSESTQSLQSLQDIINKEKKHHEVES 532  
Qy 319 KI-----GGQTVDFENK-----IRDPSCALKEKVS 345  
Db 533 QLNELKGELOTETISEHLSQSLTLAAKEKAAYATNNELSSKNSLQTLCAAFQEKLA 592  
Qy 346 DAKDLFEN-----KIGQTVDFINN-----EIRDPSCALIRKVSAGADLE 387  
Db 593 SYWQLEKNQONSSLDTSFKKLAKESHQELNNHQTTRKLDKTSLOQLQLEERAN--FE 650  
Qy 388 NKIGQTVDFINNEIR-----DPSKALIRKYTTEADLFENKIGQTVDFINKEIRDP 440  
Db 651 QK--ESTLSDENNLDLTKLLEESNKSILK--QEDVDLSLEKN-----IQTLKEDLRKS 701  
Qy 441 SKALIRKYSTEADNLE 457  
Db 702 EEARL--RFSKLEAKNLE 717

## RESULT 35

T50382

probable spindle pole body component, probable gamma-tubulin interacting protein, yeast

C:Species: Schizosaccharomyces pombe

C>Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jun-2000

C:Accession: T50382

R:Saunders, D.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, November 1999

A:Reference number: 225035

A:Accession: T50382

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-566 <SAU>

A:Cross-references: EMBL:AL133306; PIDN:CAB62095.1; GSPDB:GN00067; SPDB:SPBC902.01c

A:Experimental source: strain 972h(-); cosmid c902

C:Genetics:

A:Gene: SPBC428.20c; SPDB:SPBC902.01c

A:Map position: 2

Query Match 4.7%; Score 111.5; DB 2; Length 566;

Best Local Similarity 20.0%; Pred. No. 8.6;

Matches 101; Conservative 80; Mismatches 194; Indels 129; Gaps 23;

```

OY 4 SDSVGVTKTLAASVSVAANAAYMINSMDSYLSAVSNFAPRCSQYPKG-----56
DB 107 SSSSRDVSSTL--DEISNPINIPSTEVE-----SSNFQOTRYDQVPENPQITDMD 156
OY 57 ----SNCASAVSMRCAKQ-----DCLTLOSIXYPLEAKYQPLTLPYOL 100
DB 157 EGELESSSISIAHDSRLNRSTSTSVQHTLTREADLSSISIVLOG-----I 204
OY 101 EAAFLFKESDANPANSTEREFWMRFR--GKNHSYFHD--VFNILEKNTYADATDIE- 157
DB 205 STEYVQFKNELALSKRIPQYLLQMRALSETGLTVOELKVFSDYPSQSIDGDWVK 264
OY 158 ---NFSRRLY-MATLYTYYTNVDEFGASFNKLSTTGL-----FGWIKRAL 203
DB 265 AFINDSLQSLQSVISAEELNFALIASLQSIADASLEKPMWTIRRCIAMTQYAKL 324
OY 204 KOIIRNLPIDIGTEHSVSLQHTITSSYKDYMDQIIPALPKFAKRSIMVVOQLATVAG 263
DB 325 KLRILSSVNDNNQENKRLIOVSKYVHGPRLOELS-----DKILFEITG 373
OY 264 YVDTPWY---KKWYMKLKNFMVNRVFIPTKFFNKETREPSKALKEKYSTDTKDFENKI 320
DB 374 ----PLEYEMENWYK-----GELVDPYQEF-----VKERNGSESHD---HQ 409
OY 321 GCGTV-----DFPKETIRP--SKALKEKYSNAKDLFENKIGQYVDFINNEIRDPska 373
DB 410 GGGDVVWKGKRYFLDKELIPSLSEELVDKIFLIGKSLNFAKRGCGDFDW---AQEHYQK 465
OY 374 LIRKVTGAEDLFEKIGQGVDFINNEIRDPskALIRKYTTEADLFEK---IGQGT 429
DB 466 LVKRLSYRPHSLLEYVDKRYTESINHLV-----YLMEEVHLTDHLKAIKTYLLGQG- 519
OY 430 VDFIN-----KEIRDPskALIR 446
DB 520 -DEVLLMESLGSNDQPAVTLFR 542

```

RESULT 36

heat shock atp-dependent proteinase [imported] - Mycoplasma pulmonis (strain UAB CTRP)

C:Species: Mycoplasma pulmonis

C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001

C:Accession: G90576

R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer, I.;

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: G90576

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-842 <KUR>

A:Cross-references: GB:AL445566; PID:g14089933; PIDN:CAC13692.1; GSPDB:GN00153

A:Experimental source: strain UAB CTRP

C:Genetics:

A:Gene: MTPU\_5190

A:Genetic code: SGC3

C:Superfamily: ATP-dependent serine proteinase Ia

Query Match 4.7%; Score 111; DB 2; Length 842;

Best Local Similarity 20.6%; Pred. No. 15;

Matches 97; Conservative 73; Mismatches 150; Indels 152; Gaps 25;

```

OY 48 RIGSVPKGSNCASVSAVSMRCAKQDCLTLOSILKYPLEAKYQP--LTLPPYQLENAF- 104
DB 70 KIVAQNSKGNKIVITLLPFRKRVKNTITSETNLSKRVWVDYSPFLVADDPKRGKILYN 129
OY 105 ----ILFKESDANPANSTEREFWMRFRGKNHSYFHDVNLNLEKNYR---DADATDIE 157
DB 130 KWDIIFKNS--MPLDESKR-----NSSL-SIKYNSIAKELDKLIESDDEFL 177
OY 158 NFASRYLYMATLYKYTYTNVDEFGASFPNKLSPFTTGLFGWIKRALKOIIRSNLPIDIGT 217
DB 178 RYFSEDNNDLIFPKALN-----KIVNELS-----W-----PLELKY 210
OY 218 EHSVSLQHTITSSYKDYMD--TQIPALPKFAKRSIMVVOQLLATVAGYVDFWYKKWYM 275
DB 211 EYFIQ-----MDMINQAEFIIEFLKS-----KKEDQ 236
OY 276 KIKFMVNRVFIPTKFFNKE-----IREPSKALKEV--STDPKDLFENKIGQYVDF 328
DB 237 DLENALDEKM---KNLSKQKQEFILREKMKIRSELGSNDEEYKTKI---KDKN 288
OY 329 NKEIRDPskALKEKVSNDADLFEKIGQGVDFINNEIRDPskALIRKYSTGAED---- 384
DB 289 NKRIY--PKSVRLIANEESKLNMMSSPDANTRYIDTLMKLPKRKTSIDFLDKVA 346
OY 385 ---LFENKIG---OGTVDF---INNEIRDP--SKALIR-----KYTTEADL 421
DB 347 KILDENHYGLKEKRIIEFLAVLINNNKKNPDLKSIILKDKHDHINLVRESDDKF 406
OY 422 ENKI-----GOG-----TYDFINKEIRDPskALIRKYST 453
DB 407 DKQITNVPILALVGPPTGKTSLAIAISINKEF--VKISLGVKDAE 455

```

RESULT 37

A46417

NIP1 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: nuclear import protein; protein YW9924.01c; protein YW9952.11c; pr

C:Species: Saccharomyces cerevisiae

C>Date: 21-Sep-1993 #sequence\_revision 03-May-1994 #text\_change 06-Feb-1998

C:Accession: A46417; S53979; S59302

R:Gu, Z.; Moerschell, R.P.; Sherman, F.; Goldfarb, D.S.

Proc. Natl. Acad. Sci. U.S.A. 89, 10355-10359, 1992

A:Title: NIP1, a gene required for nuclear transport in yeast.

A:Reference number: A46417; MUID:93066237

A:Accession: A46417

A:Molecule type: DNA

A:Residues: 1-812 <GUL>

A:Cross-references: EMBL:L02899

A>Note: sequence extracted from NCBI backbone (NCBIN:117849, NCBIPI:117850)

R:Connor, R.; Churcher, C.M.

submitted to the EMBL Data Library, April 1995

A:Reference number: S53969

A:Accession: S53979

A:Molecule type: DNA

A:Residues: 571-582, 'Q', 584-640, 'K', 642, 'K', 644-812 <CON>

A:Cross-references: EMBL:249212; NID:g798940; PID:g798951; MIPS:YMR309C

R:Churcher, C.M.

submitted to the EMBL Data Library, September 1995

A:Reference number: S59302

A:Accession: S59302

A:Molecule type: DNA

A:Residues: 1-110, 'V', 112-582, 'Q', 584-602 <CHU>

A:Cross-references: EMBL:254141; NID:g1072408; PID:g984682; MIPS:YMR309C





```

Db 1897 LGEIENQLKLT-DKIDEAAVIAKKDDHSDYLCAATVSKEDWTSTFISEMLEKLP----- 1951
QY 59 CSASVASYMSRCAKQDCFLTQSLKYPLEAKYQPLTLPDPYOLEAATILFKESQANPANST 118
Db 1952 -HYMPATVVR-----LDKLPILTNDKVDKALPAADRHVATGAVTEA-----PRNDI 1998
QY 119 EKRFMMRFR-----GNHSYF-----HDLVFNLLEN 146
Db 1999 EAKLVDMEDVIGAGDIGISHHFFAGGDSIKALQIVSRSLGKLKLEKDLFANRIND 2058
QY 147 -----VTRDADATDIENFASRYLYMATLYKTYTNVDFGASF----- 184
Db 2059 LAKYVKQOSQRKNAKNTIVTGHAELEPIQKW-----YFA-----NNKEELDFHFNOSFVLFRR 2109
QY 185 -----FNKL-----SFTTGLFGM-----GIRALK 204
Db 2110 GGFDESCVAKAKNKIMEQDALRMITEKGGDFIQYNRSEFREDLFDLDYDVANGLRQAE 2169
QY 205 QI-----IRSNLPDIDGT-----EHSVSRLOHITSS-----YKDYMDTOI 239
Db 2170 KYVELATSIQKLSIRKGLVHGFIRADEGDHLLIVHHLVVDGVSWMRLFEDEFELVS 2229
QY 240 PAL-----PKFAKFSLMVYORLLATVAGVYDTPWYKWKTKLKNFVNRV 285
Db 2230 QALKGQTEIGYKTDYQOFEARLKAFAHRSRLSKAEY-----W-----FNIAKARY 2277
QY 286 -FIPTKKFFNKKEIREPSKALKEKYSTD-TKDLFENKIGQGVDFENKEIRDPKALKEKY 343
Db 2278 RFLPFPNVKLKEDYENSTLSIKLGKATADLRN-----TNRAYTEIND--ILLTAL 2329
QY 344 SDAKDLF-ENKI-----GQGVDFINN-EI-----RDPKALIRKYST 380
Db 2330 LTGARDITGENKLVKVMHEGHRDILEGYDITITIGMTTMYVLLADAGEKALSOQIKM 2389
QY 381 GAEDL--FENK-IGQGV-----DFINNE 401
Db 2390 VKETLRKIPNKGIGYGLTKYMAEDPDFTNEE 2420

```

## RESULT 40

```

HMLV
hemagglutinin precursor - Influenza A virus (strain A/PR/8/34)
C:Species: Influenza A virus
C:Date: 18-Dec-1981 #sequence-revision 18-Dec-1981 #text-change 28-May-1999
C:Accession: A93262; A90830; A04063
R:Winter G., Fields, S., Brownlee, G.G.
N:ature 292, 72-75, 1981
A:Title: Nucleotide sequence of the haemagglutinin gene of a human influenza virus HI su
A:Reference number: A93262; MUID:82013600
A:Accession: A93262
A:Molecule type: genomic RNA
A:Residues: 1-566 <WIN>
A:Cross-references: GB:V01088; GB:J02143; NID:962290; PIDN:CAA24272.1; PID:962291
R:Catton, A.J.; Brownlee, G.G.; Jewell, J.W.; Gerhard, W.
Cell 31, 417-427, 1982
A:Title: The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subu
A:Reference number: A90830; MUID:83129356
A:Contents: subtype H1
A:Accession: A90830
A:Molecule type: genomic RNA
A:Residues: 1-145, 'N', 148-155, 'E', 157-199, 'P', 201-203, 'E', 205-207, 'L', 209-268, 'W', 270-3
C:Genetics:
A:Map position: segment 4
C:Superfamily: Influenza virus hemagglutinin
C:Keywords: hemagglutinin; homotrimer; lipoprotein; thiolester bond
F:1-17/DNA: signal sequence #status predicted <SIG>
F:18-343/Product: hemagglutinin HAI chain #status predicted <HA1>
F:345-566/Product: hemagglutinin HA2 chain #status predicted <HA2>
F:555,562,565/Binding site: palmitate (Cys) (covalent) #status predicted

```

```

Query Match 4.7%; Score 110; DB 1; Length 566;
Best Local Similarity 20.8%; Pred. No. 11;
Matches 97; Conservative 67; Mismatches 171; Indels 132; Gaps 24;

```

```

QY 23 SAANAYMINSMSDY-----LSAVSDNFAERICQVPKGS-----NCSASVASYMSRCA 71
Db 101 NSFNGICITGDFIDYELERQLSSVSS--FER-ELFPKRESSPNNTTKGVTAACSHAG 157
QY 72 KOD-----CLTLQSLKYPLEAKYQPLTLPDPYOLEAATILFKESQANPANST 118
Db 158 KSSFYRNLLMLTKEGSSYP-----KLANSYVNRKKGKRYLVLMGIIHPSNSK 203
QY 119 EKRFMMRFRGKHSYFHLVFNLEKNYTRDADATDIENFASRYLYMATLYKTYTNVD 178
Db 204 DQO--NITYQEN-AVSVVTSYNNRFTPELAERKVRQAGRMNYWTLLKPGDTILF 259
QY 179 EFGASFF-NKLSFTTGL-FGMGIRALKQIIRSNLPDIDGTEHSVSRLOHITSS--YKYD 234
Db 260 EANGNLIARRYAPALSRGSG-----ITTSNASHMECNTCCQPLGAINSLSPQNI 312
QY 235 MDQIPALPKFAKFSLMVY-----QRLATVAGYDTPW--YKKWY----- 274
Db 313 HPVTIGCEPRYVSAKLRMYTGLRNIPSIQSRCLFGAIGFIEGWTGMIDGWYGHQNI 372
QY 275 -----MKLKNFVNRV-----FIPTKKFFNKKEIREPSKALKEKYSTD 311
Db 373 EOGSGYADQKSTQNAINGITRNKVSVEIEKNIQFTAVGKEFNK-DEKREMNKKRYVDG 431
QY 312 TKD-----LFPNKIGQGVDFENKEIRDPKALKEKYSTDAKDLFENKIGQGVTD 361
Db 432 FLDIWTYNNMELLVLEN--ERTLDPHDSNVKMLYEYVVSQLNNAKE-----IGNGCFE 483
QY 362 F--INNEIRDPKALIRKYSTGAEDL--FENKIGQGVDFINNE 401
Db 484 FYHKCDNE-----CMESYRNGTYDYPKYSSESKLNREKRVQVKLE 523

```

Search completed: October 19, 2002, 07:52:25  
 Job time : 55 secs

**THIS PAGE BLANK (USPTO)**